PB89-118483

Relating Error Bounds for Maximum Concentration Estimates to Diffusion Meteorology Uncertainty (Journal Version)

(U.S.) Environmental Protection Agency
Research Triangle Park, NC

1987

U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before complet

1. REPORT NO. EPA/600/J-87/447	2.	3. PB89-118483
4. TITLE AND SUBTITLE: RELATING ERROR BOUNDS FOR MAXIMUM CONCENTRATION ESTIMATES TO DIFFUSION METEOROLOGY UNCERTAINTY (Journal Version)		5. REPORT DATE: 1987 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S): J.S. Irwin, S.T. Rao*, W.B. Petersen, and D.B. Turner		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS: EPA EOB/MD/ASRL RTP, NC 27711 *Division of Air Resources N.Y. Dept. Environ. Conservation Albany, NY		10. PROGRAM ELEMENT NO. CDTA1D/04 - 6240 (FY-86) 11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS: Atmospheric Sciences Research Laboratory -- RTP, NC Office of Research and Development U. S. Environmental Protection Agency Research Triangle Park, NC 27711		13. TYPE OF REPORT AND PERIOD COVERED: Journal Article 14. SPONSORING AGENCY CODE: EPA/600/09

15. SUPPLEMENTARY NOTES: Atmospheric Environment, 21(9):1927-1937, 1987.

16. ABSTRACT

This paper relates the magnitude of the error bounds of data, used as inputs to a Gaussian dispersion model, to the magnitude of the error bounds of the model output. The research addresses the uncertainty in estimating the maximum concentrations from elevated buoyant sources during unstable atmospheric conditions, as these are most often of practical concern in regulatory decision making. The ability to develop specific error bounds, tailored to the modeling situation, allows more informed application of the model estimates to the air quality issues. The numerical uncertainty analysis is performed using the Monte-Carlo technique to propagate the uncertainties associated with the model input. Uncertainties were assumed to exist in four model input parameters: 1) wind speed, 2) standard deviation of lateral wind direction fluctuations, 3) standard deviation of vertical wind direction fluctuations, and 4) plume rise. The authors conclude that the error bounds for the estimated maximum concentration and the distance to the maximum can be double that of the error bounds for individual model input parameters. These results allow estimation of minimum bounds on errors in model output when considering reasonable input error bounds.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL INFORMATION SERVICE
SPRINGFIELD, VA. 22161

18. DISTRIBUTION STATEMENT: RELEASE TO PUBLIC	19. SECURITY CLASS (This Report): UNCLASSIFIED 20. SECURITY CLASS (This page): UNCLASSIFIED	21. NO. OF PAGES: 13 22. PRICE

EPA/600/J-87/447
JOURNAL ARTICLE

Atmospheric Environment Vol. 21, No. 9, pp. 1927–1937, 1987.
Printed in Great Britain.

Pergamon Journals Ltd.

NTIS is authorized to reproduce and sell this report. Permission for further reproduction must be obtained from the copyright owner.

RELATING ERROR BOUNDS FOR MAXIMUM CONCENTRATION ESTIMATES TO DIFFUSION METEOROLOGY UNCERTAINTY

J. S. IRWIN*, S. T. RAO†, W. B. PETERSEN* and D. B. TURNER*‡

*Meteorology and Assessment Division, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, U.S.A. (on assignment from the National Oceanic & Atmospheric Administration) and †Division of Air Resources, New York State Department of Environmental Conservation, Albany, NY 12233, U.S.A.

(*First received* 11 *August* 1986 *and in final form* 26 *January* 1987)

Abstract—The intent of this paper is to relate the magnitude of the error bounds of data, used as inputs to a Gaussian dispersion model, to the magnitude of the error bounds of the model output, which include the estimates of the maximum concentration and the distance to that maximum. The research specifically addresses the uncertainty in estimating the maximum concentrations from elevated buoyant sources during unstable atmospheric conditions, as these are most often of practical concern in regulatory decision making. A direct and quantitative link between the nature and magnitude of the input uncertainty and modeling results has not been previously investigated extensively. The ability to develop specific error bounds, tailored to the modeling situation, allows more informed application of the model estimates to the air quality issues.

In this study, a numerical uncertainty analysis is performed using the Monte-Carlo technique to propagate the uncertainties associated with the model input. Uncertainties were assumed to exist in four model input parameters: (1) wind speed, (2) standard deviation of lateral wind direction fluctuations, (3) standard deviation of vertical wind direction fluctuations, and (4) plume rise. For each simulation, results were summarized characterizing the uncertainty in four features of the ground-level concentration pattern predicted by the model: (1) the magnitude of the maximum concentration, (2) the distance to the maximum concentration, and (3) and (4) the areas enclosed within the isopleths of 50% and 25% of the error-free estimate of maximum concentration.

The authors conclude that the error bounds for the estimated maximum concentration and the distance to the maximum can be double that of the error bounds for individual model input parameters. The model output error bounds for the areas enclosed within isopleth values can be triple the error bounds of the input. It was not our intent to cover all possible combinations for the error in the input parameters. Ours was a much more limited goal of providing a *lower* bound estimate of model uncertainty in which we assume the input is reasonably well characterized and there is no bias in the input. These results allow estimation of *minimum* bounds on errors in model output when considering reasonable input error bounds.

Key word index: Error bounds, maximum concentration, uncertainty, diffusion meteorology, Gaussian models, Monte-Carlo sampling.

1. INTRODUCTION

There are several sources of uncertainty in all simulations of air quality. One is the inherent lack of predictability of transport and diffusion imposed by the stochastic nature of the atmosphere. Even perfect knowledge of the initial state of the atmosphere would not allow reduction of the imprecision below a fundamental level. A second source of uncertainty arises because the initial state of the atmosphere is always imperfectly known. Although the input uncertainty can be due to a number of causes (errors in measurement, errors in estimation, non-representativeness), we are not concerned in this study with the origin of the uncertainty in the inputs, but with its magnitude regardless of cause. A direct and quantitative link between the nature and magnitude of the input uncertainty and the modeling results has not been previously investigated extensively.

Intuitively, one expects the output of a model to be at least as uncertain, and perhaps more so, than those associated with the inputs to the model. One might also expect that the distribution of the modeled values may be similar to the distribution associated with the input errors. These notions, relating input and model uncertainty, are reasonable if the modeled result is a linear function of the input values. Our goal in this study is to investigate whether this view is reasonable for Gaussian-plume modeling of diffusion from an elevated point source. We also seek to quantify the model uncertainty in terms of the assumed input uncertainty.

For this study, we used a Monte-Carlo method to propagate the uncertainties associated with four specific model inputs (wind speed, plume rise, and the standard deviations of azimuth and elevation angles of the wind). This involved iteratively selecting values for each model input from their respective probability density functions in a random fashion. A Gaussian-plume diffusion model then provided estimates of ground-level concentration associated with each input

‡To whom correspondence should be addressed.

parameter set. Each simulation consisted of 10,000 random value parameter sets.

The intent is to quantitatively characterize the error bounds for concentration estimates from a Gaussian-plume model resulting from model input uncertainty. The research specifically addresses the uncertainty of maximum concentration values as these are most often of practical concern in regulatory decision making. Because of its importance to the occurrence of maximum concentrations from medium to large buoyant sources, an unstable atmospheric condition, typical of sunny afternoons, is the case selected for examination. The ability to develop specific error bounds, tailored to the modeling situation, allows more informed application of the model estimates to the air quality issues.

Near the completion of the work reported here, an article was published by Freeman *et al.* (1986). Although that paper has similar goals and also uses Monte-Carlo techniques, details are different as discussed later in this paper.

2. ANALYSIS AND PROCEDURES

In the following, we discuss the uncertainty in short-term (hourly) ground-level concentration estimates, as might be modeled from a large isolated point-source. Concentrations were estimated using a Gaussian-plume dispersion model. Gaussian dispersion modeling is most applicable for elevated releases within the near-neutral boundary layer (Gryning *et al.*, 1987). Model estimates for use in regulatory decision making most often focus concern on the location and magnitude of the maximum concentration. Hence, we investigated the uncertainty in the following features of the ground-level concentration pattern:

- Cm the maximum concentration,
- Dm the distance to the maximum concentration,
- $A50$ the area coverage having concentration values greater than $Cm^{\circ}/2$,
- $A25$ the area coverage having concentrations values greater than $Cm^{\circ}/4$,

where Cm° is the maximum concentration resulting for the initial, error-free, input data to the dispersion model. Throughout this discussion 'error-free' refers to results obtained if there were no uncertainty in the input values. The first two features address questions of how precise a modeling result is for the peak impact, given input data with particular accuracy. The latter two features address questions of the adequacy of a receptor network design for observing an impact.

Describing input uncertainty

Having chosen the features of the ground-level concentration pattern to study, our next concern was to develop a reasonable characterization for the uncertainty in the meteorological input. An obvious choice would be to assume that the errors in model input are Gaussian in nature, as in Freeman *et al.* (1986). However, there is reason to believe that the errors, being bounded below by zero, would have a skewed distribution. An attempt to properly approximate the distribution of errors was made by comparing estimates from a meteorological processor with measured data on a tower. Details of this are given in Appendix A. It was concluded that a log-normal distribution of the ratios (value with error to error-free value) was more appropriate for characterizing the uncertainties in the model input.

In the succeeding analyses, the uncertainty of dispersion model input is characterized using a log-normal distribution of the ratios. We defined the distributions by the median, M, and the 75th percentile value, $P75$. Our intent was to study the effects of input uncertainty on model output. The effects of bias in the model input can to a first order be treated separately and have been discussed previously in sensitivity studies, e.g. Keyser and Anthes (1976) and Hilst (1970). Therefore, for our simulations, the median was assumed to be 1.0 (no bias).

The inverse of the $P75$ value in the ratio cumulative frequency distribution (*cfd*), when there is no bias, is the $P25$ value where the distribution is log-normal; thus the $P75$ value will identify the limits within which are 50% of the values, the interquartile range. For instance, if 50% of the Cm values are between $1.5Cm^{\circ}$ and $(1/1.5)Cm^{\circ}$ and the uncertainty is unbiased; this is the same as defining the median of the Cm/Cm° distribution to be 1.0 and the $P75$ value to be 1.5.

Gaussian-plume model and characteristics calculated

Ground-level concentrations at the downwind distance, X, beneath the plume centerline from an elevated source were calculated using,

$$C = Q/(\pi U\, \sigma_y \sigma_z).\ \mathrm{Exp}\,[-0.5(H/\sigma_z)^2]$$

where Q is emission rate ($g\,s^{-1}$), U is wind speed at the stack height ($m\,s^{-1}$), H is effective plume height (m), and σ_y and σ_z are the horizontal and vertical standard deviations (m), of plume concentration distribution at the downwind distance X. The effective plume height H is $h + \Delta h$, where h is the physical stack height and Δh is the final plume rise. σ_z and σ_y were estimated using,

$$\sigma_{z,y} = \sigma_{e,a} X f_{z,y},$$

where σ_e and σ_a are the standard deviations of the vertical and lateral wind direction fluctuations in radians and,

$$f_{z,y} = 1/[1 + 0.9(t/t_{z,y})^{1/2}]$$

where $t = X/U$ and t_z is 500 s (unstable) and 50 s (stable), and t_y is 1000 s. This is Model 4 of Irwin (1983) and is a synthesis of schemes suggested by Draxler (1976) for modeling the dispersion. Calculations presented are for unstable conditions only.

Effects of buoyancy-induced dispersion were included in these calculations following Pasquill's (1976) procedures by substituting effective values for σ_y and

σ_z in the above equations where,

$$\sigma_{ye} = [\sigma_y^2 + (\Delta h)^2/10]^{1/2}$$

and $$\sigma_{ze} = [\sigma_z^2 + (\Delta h)^2/10]^{1/2}.$$

For the calculations presented, we modeled the terrain as flat with no depositional or chemical loss. We did not include the effects of limited mixing in the vertical due to the presence of an elevated temperature inversion. Limited mixing will not normally affect maximum concentrations.

The four parameters Δh, U, σ_a and σ_e were assumed to contain uncertainty. Many of the calculations employed a 'base case' set of median values for the model input: $U = 3\ \mathrm{m\,s^{-1}}$, $\sigma_a = 26°$ and $\sigma_e = 16°$. Q and h were assumed to be error free. We normalized the results for variations in Q using $Q = 1$. A series of values for h and Δh were used; this is discussed in Appendix B.

Values of the four characteristics were determined for each simulation: the maximum concentration, the distance to the maximum concentration and approximations to the area included within two isopleths of equal concentration. The isopleth values were for concentration equal to $0.5Cm°$ (we called this area $A50$) and for concentration equal to $0.25Cm°$ ($A25$), where $Cm°$ is the maximum concentration computed with 'error-free' input.

The maximum concentration and the distance to the maximum were determined by a numerical search from solutions at various distances. To approximate the area within a concentration isopleth, the near and far positions of the isopleth, $X1$ and $X3$ in Fig. 1, were determined iteratively, similar to the search performed to determine the value and position of the maximum ($X2$ in Fig. 1). The isopleth half-width $Y2$ at the distance for the maximum is determined from,

$$Y2 = \sigma_{ye}[2\ln(Cm°/Ci)]^{0.5}$$

where Ci is either $0.5Cm°$ or $0.25Cm°$.

By approximating the isopleth as an ellipse, we can estimate the enclosed area. The area of an ellipse is πab, where a is half the major axis of the ellipse and b is half the minor axis (see Fig. 1). Tests determined that defining a and b as

$$a = (X3 - X1)/2$$

$$b^2 = Y2^2/\{1 - [a - (X2 - X1)]^2/a^2\}$$

approximated the area determined by numerical integration to within 15%.

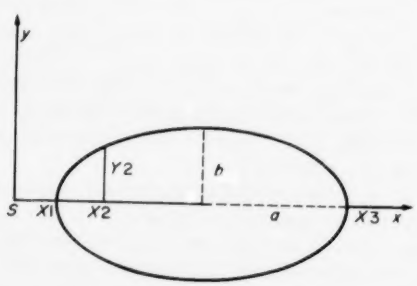

Fig. 1. Schematic of area computation. S is the source location (assumed to be at the origin of the x–y coordinate system) and $X2$ is the distance to the maximum concentration. $X1$, $X3$ and $Y2$ are the positions determined along and off axis of the plume centerline of the isopleth of concern. a and b are half the major and minor axes of the elipse.

Monte-Carlo simulation procedure

Of the available error propagation methods, we selected the Monte-Carlo method. It presents a simple and direct approach to the problem and is appropriate for use when the errors are independent. This approach enables an evaluation of the output of the model for many sets of combinations of the input parameters. We obtain these combinations of input values by random sampling from the distribution assigned to each of the four input parameters.

The selection is independent for each variable in the input variable set. With each input variable set, we model the surface concentrations and derive the output variables of interest, namely Cm, Dm, $A50$ and $A25$. We selected 10,000 input variable sets. Thus, the output of the Monte-Carlo process consists of 10,000 values for each of the output variables considered in this study. The 10,000 values for each output variable represent a characterization of the uncertainty in that feature of the concentration pattern, for instance, maximum concentration or distance to the maximum.

This characterization is an estimate, based on the sample drawn. The precision of the characterization is sensitive to the sample size. The procedure is basically a simple random selection process for developing a sample. The sample is a characterization of the uncertainty in the modeling results. In an effort to study the sensitvity of the simulation results to sample size, we performed a case study for sample sizes of 100, 500, 1000, 5000 and 10,000. Ten simulations were run at each sample size. The variation in the computed interquartile range for the maximum concentration distributions reduced from 40% to less than 5%, in going from use of 100 samples to 10,000 samples in the Monte-Carlo method. Hence, a sample size of 10,000 was considered adequate for use in the uncertainty analyses.

One of the potential limitations of the Monte-Carlo simulation is its cost in computer simulation time. This was a minor concern in this study as the simulations were performed on a personal computer; thus there was no charge per block of time used.

Another important limitation results from assigning the uncertainty distributions to each of the input variables. We minimized the impact of this by selecting distributions similar to those observed to occur. A log-normal distribution of ratios seemed to be a realistic choice from examination of data from Savannah River Laboratory. A description of the data collected is given

in Crawford (1974) (see Appendix A). In reality, the distributions may differ from a lognormal distribution. It may be that the uncertainty distribution of the plume rise is more Gaussian. We ran several simulations using a Gaussian distribution to characterize the plume rise uncertainty. The results were indistinguishable to those developed using a log-normal distribution. This is encouraging but in no way eliminates the restrictions inherent to the results presented.

Our use of log-normal distributions of the ratios differs from that of Freeman *et al.* (1986) who used Gaussian distributions for the errors in the input parameters. It was probably because of the assumed Gaussian distribution that Freeman *et al.* found it necessary to remove outliers. In our use of the log-normal distributions of ratios it was not necessary to consider any outliers; each generated data input was usable.

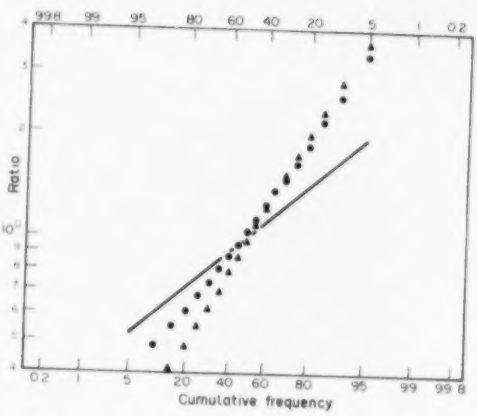

Fig. 3. Same as Fig. 3, except $P75$ equals 1.3.

3. RESULTS

Each simulation of 10,000 sets of calculations yielded a value for the four characteristics which was expressed as a ratio (value over error free value). A separate file of 10,000 values for each characteristic was ordered by sorting so that a cumulative frequency distribution was obtained. Ratio values for specific percentiles were then printed. These values could then be plotted. Figure 2 is a log probability plot typical of the results obtained. For these results, the $P75$ value was 1.10 for each of the input parameters, the source height was 200 m and the plume rise was 400 m. Figure 3 summarizes the results for the same source and meteorology as in Fig. 2 except for the $P75$ value which was 1.3 to define the uncertainty in the model input. Although the source height and plume rise are typical of many sources, a $P75$ value of 1.10 in the meteorological input is unrealistically small compared to input estimation capabilities. In our comparisons with the SRL tower data, $P75$ values on the order of 1.3–1.4 occurred frequently. Hence, we anticipate that for routine situations, $P75$ of 1.2–1.3 is a conservative estimate of the input uncertainty from routinely available wind speed and turbulence values. With research grade meteorological instrumentation on site, a $P75$ of 1.1–1.2 might be anticipated.

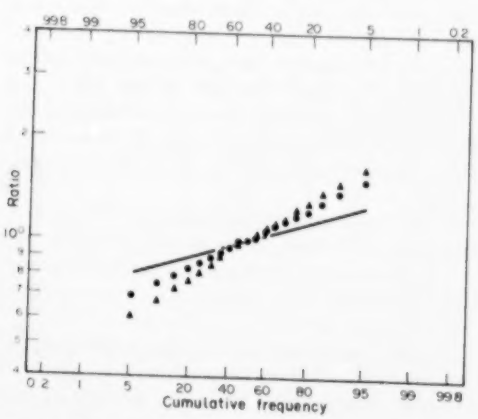

Fig. 2. Distribution of ratios of concentration maximum (triangles) and distance to maximum (dots) for the case of input uncertainty having $P75$ equal to 1.1. The solid line is the input uncertainty distribution.

Characterization of results

We conducted a goodness of fit analysis on the *cfd* depicted in Fig. 2, following the procedures outlined by Holland and Fitz-Simons (1982). Seven distributions were tested: normal, Gamma, Weibull, Beta and three log-normal distributions. The log-normal distributions were the 2-parameter, 3-parameter and 4-parameter (Johnson SB distribution). We could not reject the 2-parameter log-normal distribution as a possible fit to these data, using a Kolmogorov–Smirnov test at the 1% significance level. Of course, the 3- and 4-parameter log-normal distributions also fit the data, but we considered these unduly complex for our purposes of providing a simple methodology to summarize the uncertainty results. For simulations with $P75$ less than 1.3, the distributions of model output were consistently well characterized by log-normal distributions. However, as can be seen in Fig. 3, for $P75$ of 1.3, there is some departure from a log-normal fit.

Our objective was to develop a robust characterization of model uncertainty, typical for routine modeling studies. It was not our intent to simulate all possible combinations of input error combinations in which the $P75$ values would no longer be the same for each model input. Ours was a much more limited goal, of providing a lower bound estimate of the model uncertainty for cases in which the input is reasonably well characterized ($P75$ is between 1.1 and 1.3) and there is no bias in the input. With such, we can specify the smallest error bounds to be expected. This is quite

useful as it specifies the minimum error bound limits to be expected under optimum conditions.

Tables 1 and 2 summarize the quartile values determined for the basic set of computations conducted (for three stack heights and three plume rise values). In general, as the size of the error in the input increases, the median value for the ratio deviates from unity. Also, the variation in the median value for the areas of the two isopleths is much greater than that for the maximum concentration. This variation is larger than that seen for the medians determined for the distance to the maximum concentration. Also, the distribution of the ratio values for the areas deviates from log-normality more so than for the other features.

Other special cases were conducted to address the sensitivity of the results to the values chosen for the true wind speed and the turbulence intensities. These latter cases revealed that all other factors being equal, multiplying the base case wind speed by two, or by four, resulted in no change in the distribution of the resulting concentration ratios. Likewise, multiplying the base case turbulence values by 0.5, or two, resulted in no change in the distribution of the concentration ratios.

Figures 4 and 5 summarize the results obtained for the quartiles (the 25th, 50th and 75th percentile values), for the distance to maximum and maximum concentration ratio distributions, when the stack height is 200 m and the plume rise is 400 m. The interquartile range, IQR, is equal to the 75th percentile value minus the 25th percentile value.

Summarizing the results

For practical application of the above results, we sought a means to summarize the cumulative frequency distributions of the four ratios. We wanted to do this in terms of the dispersion model input values and the uncertainty assumed in these values. Our solution was to quantify the variation of the output interquartile range, IQR, and the output median, M, of these distributions as functions of the $P75$ value.

Within Appendix B, we present characterizations of M and IQR that account for the lesser effects of stack height and plume rise. For most purposes, it will be sufficient to estimate IQR and M as functions of $P75$ as:

$$IQR = A(P75 - 1) + B \tag{1a}$$

$$M = C(P75 - 1) + D. \tag{1b}$$

Table 3 lists the constants A, B, C and D for the four ratio distributions.

Table 1. Quartile values for simulation results of the maximum concentration and the distance to the maximum

		Maximum concentration ratio quartile values								
		$P75 = 1.1$			$P75 = 1.2$			$P75 = 1.3$		
h	Δh	$Q25$	$Q50$	$Q75$	$Q25$	$Q50$	$Q75$	$Q25$	$Q50$	$Q75$
100	100	0.83	0.99	1.20	0.69	0.99	1.41	0.57	0.97	1.62
	400	0.80	1.00	1.25	0.64	0.98	1.50	0.52	0.97	1.81
	700	0.79	1.00	1.25	0.62	0.98	1.52	0.52	1.00	1.85
200	100	0.83	1.00	1.19	0.70	0.99	1.38	0.59	0.98	1.58
	400	0.81	1.00	1.23	0.66	0.99	1.46	0.55	0.97	1.70
	700	0.80	1.00	1.24	0.64	1.00	1.50	0.53	0.97	1.78
300	100	0.84	0.99	1.17	0.71	0.99	1.36	0.60	0.97	1.54
	400	0.82	0.99	1.21	0.68	0.99	1.44	0.57	0.99	1.67
	700	0.81	1.00	1.23	0.66	0.99	1.45	0.54	0.97	1.71

		Distance to maximum ratio quartile values								
		$P75 = 1.1$			$P75 = 1.2$			$P75 = 1.3$		
h	Δh	$Q25$	$Q50$	$Q75$	$Q25$	$Q50$	$Q75$	$Q25$	$Q50$	$Q75$
100	100	0.88	1.01	1.15	0.80	1.01	1.31	0.73	1.03	1.50
	400	0.85	1.00	1.19	0.75	1.01	1.39	0.67	1.03	1.66
	700	0.84	1.00	1.21	0.73	1.03	1.46	0.64	1.03	1.72
200	100	0.88	1.00	1.14	0.78	1.00	1.30	0.71	1.02	1.49
	400	0.85	1.00	1.18	0.75	1.02	1.40	0.68	1.03	1.62
	700	0.84	1.00	1.20	0.73	1.01	1.43	0.64	1.02	1.70
300	100	0.88	1.01	1.16	0.79	1.01	1.32	0.70	1.02	1.49
	400	0.86	1.01	1.19	0.75	1.01	1.37	0.68	1.02	1.59
	700	0.84	1.00	1.19	0.73	1.01	1.42	0.64	1.03	1.69

Results are tabulated as a function of stack height, h, plume rise, Δh, and 75th percentile value, $P75$, used to characterize the uncertainty to be associated with the wind speed and turbulence values. The central wind speed was 3.0 m s^{-1}, σ_a was 26 degrees, and σ_e was 16 degrees.

Table 2. Quartile values for simulation results of the area within the isopleths of 50% and 25% of the maximum concentration

		Area of 50% isopleth ratio quartile values								
		$P75 = 1.1$			$P75 = 1.2$			$P75 = 1.3$		
h	Δh	$Q25$	$Q50$	$Q75$	$Q25$	$Q50$	$Q75$	$Q25$	$Q50$	$Q75$
100	100	0.70	0.96	1.27	0.42	0.87	1.50	0.15	0.76	1.68
	400	0.65	0.96	1.32	0.32	0.83	1.54	0.06	0.73	1.83
	700	0.63	0.96	1.33	0.29	0.85	1.63	0.05	0.74	1.96
200	100	0.71	0.96	1.27	0.42	0.87	1.49	0.20	0.76	1.71
	400	0.66	0.96	1.32	0.36	0.86	1.60	0.12	0.74	1.78
	700	0.63	0.95	1.34	0.31	0.84	1.63	0.07	0.71	1.90
300	100	0.70	0.97	1.29	0.44	0.87	1.52	0.20	0.75	1.74
	400	0.66	0.96	1.32	0.37	0.85	1.58	0.14	0.75	1.86
	700	0.65	0.96	1.33	0.34	0.83	1.58	0.09	0.72	1.91

		Area of 25% isopleth ratio quartile values								
		$P75 = 1.1$			$P75 = 1.2$			$P75 = 1.3$		
h	Δh	$Q25$	$Q50$	$Q75$	$Q25$	$Q50$	$Q75$	$Q25$	$Q50$	$Q75$
100	100	0.80	0.99	1.23	0.62	0.94	1.44	0.44	0.86	1.63
	400	0.76	0.97	1.26	0.54	0.90	1.48	0.38	0.83	1.76
	700	0.75	0.98	1.27	0.53	0.92	1.55	0.35	0.83	1.86
200	100	0.78	0.99	1.24	0.59	0.93	1.47	0.44	0.88	1.72
	400	0.76	0.98	1.27	0.56	0.93	1.56	0.38	0.86	1.77
	700	0.74	0.98	1.29	0.53	0.92	1.57	0.35	0.84	1.88
300	100	0.77	0.99	1.25	0.60	0.94	1.52	0.42	0.86	1.75
	400	0.76	0.98	1.28	0.55	0.92	1.54	0.40	0.86	1.85
	700	0.75	0.99	1.29	0.54	0.91	1.55	0.37	0.87	1.86

Results are tabulated as a function of stack height, h, plume rise, Δh, and 75th percentile value, $P75$, used to characterize the uncertainty to be associated with the wind speed and turbulence values. The central wind speed was 3.0 m s^{-1}, σ_a was 26 degrees, and σ_e was 16 degrees.

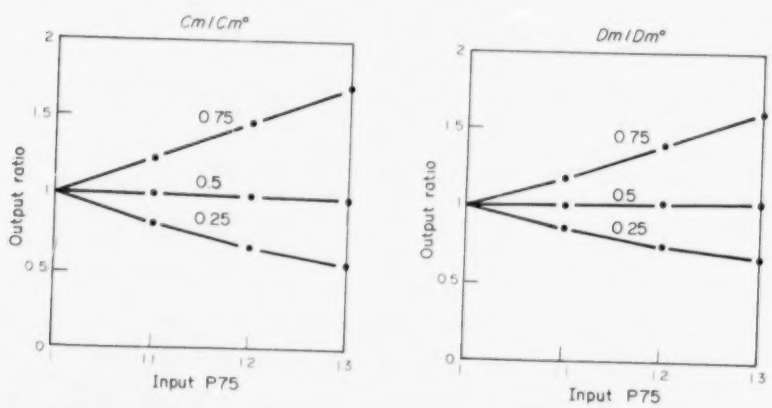

Fig. 4. Variation of the 75th and 25th percentiles and median for the distributions of ratios of maximum concentration and distance to maximum, as a function of input uncertainty.

The regression results presented in Table 3 provide a means for constructing interquartile error bounds to describe the uncertainty of the modeling results. Typically, the 50% or 95% confidence bounds are discussed in the literature. Of course, one could construct error bounds for any probability. This is demonstrated in Appendix C. Using the methodology presented in Appendix C the error bounds for six different percentiles were computed. (Results are available from the authors.) The conclusions stated here for the 50% error bounds are generally true for error bounds for other percentiles.

Results can be summarized by the following two simple 'rules of thumb.' Considering the input errors expressed as $P75$ between 1.05 and 1.3, the output ratio, $C^{\circ}(p)$, in the cumulative frequency distribution is

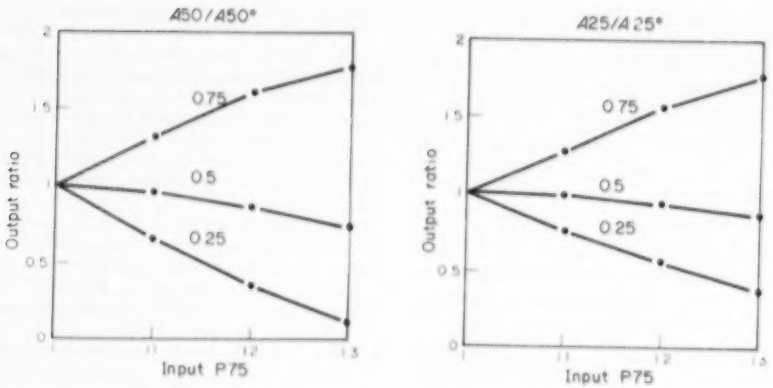

Fig. 5. Same as Fig. 4, except for the distribution of ratios of area of 50% and 25% isopleths of maximum concentration.

Table 3. Linear regression results for the interquartile range, *IQR*, and median, *M*, of the simulated distributions of modeling results

	IQR	*A*	*B*	r^2
Cm/*Cm*°		3.68	0.04	0.92
Dm/*Dm*°		3.04	0.02	0.90
*A*50/*A*50°		5.28	0.13	0.95
*A*25/*A*25°		4.48	0.06	0.96
	M	*C*	*D*	r^2
Cm/*Cm*°		−0.10	1.01	0.62
Dm/*Dm*°		0.11	0.99	0.73
*A*50/*A*50°		−1.09	1.07	0.98
*A*25/*A*25°		−0.63	1.05	0.94

A, *B*, *C* and *D* are the constants for use in Equations (1a) and (1b). *r* is the Pearson correlation coefficients.

approximately related to the input ratio for the same percentile, $C^i(p)$, by

$$C^o(p) \approx 1 + 2[C^i(p) - 1]$$

for the maximum concentration and the distance to maximum and by

$$C^o(p) \approx 1 + 3[C^i(p) - 1]$$

for the *A*50 and *A*25 isopleth areas.

The above rules of thumb reveal the variability of the modeling results to input uncertainty. Since there may be interest in other than the results for the interquartile range, information for other *Z* (the reduced normal variate) values for each of the four features for three input error bounds are available from the authors.

4. CONCLUSIONS

For meteorological conditions that frequently cause maximum concentrations from medium to large buoyant sources (unstable), a magnitude of the error bounds on four features of model output have been related to the magnitude of the error bounds of four model input parameters, three meteorological parameters: wind speed, and standard deviations of both horizontal and vertical wind angles, and one parameter related to meteorology: plume rise. For error distributions of the input parameters, we assumed that the ratios (value with error to error-free value) were log-normal. For the investigation here the error bounds of those ratios were assumed equal for the four input parameters.

The rules of thumb that resulted from the analysis are: for the magnitude of the maximum concentration and for the distance to the maximum concentration, the error bounds for these features are approximately double the error bounds for the individual input parameters. For the areas within the isopleths of 50% and 25% of the error free maximum the error bounds are approximately three times those of the individual input parameters. Although emphasis was placed upon the 50% error bounds in the discussion, analysis of other error bound ranges (up to the 95% error bounds) indicated that the above rules of thumb are still valid.

Although one could investigate ad infinitum other meteorological conditions and various combinations of assumptions using different magnitudes of error bounds for the input parameters, it was not our intent to exhaustively examine this complex area. Rather, we consider the study successful in finding the above rules of thumb for this simple meteorological condition, for the situation of equal error bounds for the four parameters, and in showing that Monte-Carlo techniques are useful to explore the influence of input errors upon model output.

These results allow estimation of minimum bounds (minimum because we have not considered all sources of uncertainty) on output errors when considering reasonable input error bounds. An estimate of the magnitude of the 50% error bound for measurements from well sited research grade instrumentation is on

the order of 10% of the mean. The estimate of the 50% error bound on the input parameters is more likely to be 20% of the mean value when one considers siting difficulties and problems of interpolation or extrapolation to plume height. If we assume 20% error, this translates to a 40% error in the magnitude of the maximum concentration estimates by Gaussian air quality models.

Input uncertainties can mask poor and good model performance alike. A knowledge of the effects of input error bounds provides a basis for developing reasonable margins of safety. It also provides a basis for gauging how discerning model evaluation can be.

Future research should focus on better defining the nature and character of diffusion model input uncertainties. Such studies allow refinement in our understanding of the inherent limitations imposed by model input uncertainties. The Monte-Carlo method presented here is general and flexible. It is applicable for use with different diffusion models or different models of input uncertainty. As applied here, the technique enables one to systematically evaluate the overall uncertainty in the modeling results.

Acknowledgements—The authors thank Drs Cynthia Hirtzel and Gopal Sistla, and Jia-Yeong Ku for many helpful discussions on the subject and their assistance in the preparation of this paper; and Sylvia Coltrane for her secretarial assistance.

REFERENCES

Benkley C. W. and Schulman L. L. (1979) Estimating hourly mixing depths from historical meteorological data. *J. appl. Met.* **18,** 772–780.

Crawford T. V. (1974) Progress report—dose-to-man program FY 1973. USAEC Report DP-1341, E. I. duPont deNemours and Co. Savannah River Laboratory, Aiken, SC.

Deardorff J. W. (1974) Three-dimensional numerical study of the height and mean structure of a heated planetary boundary layer. *Boundary-Layer Met.* **7,** 81–106.

Draxler R. R. (1976) Determination of atmospheric diffusion parameters. *Atmospheric Environment* **10,** 99–105.

Freeman D. L., Egami R. T., Robinson N. F. and Watson J. G. (1986) A method for propagating measurement uncertainties through dispersion models. *J. Air Pollut. Control. Ass.* **36,** 246–253.

Garrett A. J. (1981) Comparison of observed mixed-layer depths to model estimates using observed temperatures and winds, and MOS forecast. *J. appl. Met.* **11,** 1277–1283.

Gryning S. E., Holtslag A. A. M. Irwin J. S. and Sivertsen B. (1987) Applied dispersion modelling based on meteorological scaling parameters. *Atmospheric Environment* **21,** 79–89.

Hilst G. R. (1970) Sensitivities of air quality prediction to input errors and uncertainties. *Proceedings of Symposium On Multiple-Source Urban Diffusion Models* (edited by Stern A. C.). AP-86, Research Triangle Park, NC 27711.

Holland D. M. and Fitz-Simons T. (1982) Fitting statistical distributions to air quality data by the maximum likelihood method. *Atmospheric Environment* **16,** 1071–10[illegible].

Holtslag A. A. M. (1984) Estimates of diabatic wind speed profiles from near surface weather observations. *Boundary-Layer Met.* **29,** 225–250.

Holtslag A. A. M., Gryning S. E., Irwin J. S. and Sivertsen B. (1985) Parameterization of the atmospheric boundary layer for air pollution dispersion models. In *Air Pollution Modeling and Its Application* (edited by De Wispelaere V. C.). Plenum Press, New York.

Irwin J. S. (1983) Estimating plume dispersion—a comparison of sigma schemes. *J. Clim. appl. Met.* **22,** 92–114.

Irwin J. S. (1985) Modeling the diurnal variation of meteorological variability within the boundary layer—preliminary comparison results. In *Air Pollution Modeling and Its Application* (edited by De Wispelaere V. C.). Plenum Press, New York.

Keyser D. and Anthes R. A. (1976) Sensitivity tests with a parameterized mixed-layer model suitable for air quality simulations. *Proceedings of Conference on Environmental Modeling and Simulation.* EPA 600/9-76-016. U.S. Environmental Protection Agency, Research Triangle Park, NC, 27711, pp. 313–317.

Pasquill F. (1976) Atmospheric dispersion parameters in Gaussian plume modeling—II. Possible requirements for change in the Turner workbook values. EPA 600/4-76-030b. U.S. Environmental Protection Agency, Research Triangle Park, NC.

Paumier J. P., Stinson D., Kelly T., Bollinger C. and Irwin J. S. (1986) MPDA-1.0—a meteorological processor for diffusion analysis. EPA 600/8-86/011. U.S. Environmental Protection Agency, Research Triangle Park, NC.

Tennekes H. (1973) A model for the dynamics of the inversion above a convective boundary layer. *J. Atmos. Sci.*, **30,** 558–567.

Van Dop H., de Haan B. J. and Cats G. J. (1981) Meteorological input for a three dimensional medium range air quality model. In *Air Pollution Modeling and Its Application. I* (edited by De Wispelaere V. C.), pp. 181–190. Plenum Press, New York.

APPENDIX A. ESTIMATING INPUT DISTRIBUTIONS

To better understand the nature of the uncertainty in meteorological input, we compared wind and turbulence values simulated by a meteorological processor with those observed on a tall meteorological tower. The meteorological simulation results and tower data were available from another study (Irwin, 1985).

The tower measurements were collected during June 1976 by the Meteorology Group of the Savannah River Laboratory (SRL), Atomic Energy Division, E. I. du Pont de Nemours & Company. These data contain 15-min averages and standard deviations of temperature, wind speed, horizontal wind direction and vertical wind direction (Crawford, 1974). The data were collected on the WJBF tower at seven elevations above ground: 10, 36, 91, 137, 182, 243 and 304 m. The WJBF—SRL Meteorological Facility, is 11 km southeast of Augusta, Georgia and 23 km southwest of Aiken, South Carolina.

The estimates of wind and turbulence were developed using a meteorological processor (Irwin, 1985; Paumier *et al.*, 1986). The meteorological processor offers a simplistic set of methods for processing three types of meteorological data for use in dispersion analyses: the National Weather Service (NWS) hourly weather observations, the NWS twice-daily upper air reports and user supplied onsite data. The processor offers a selection of methods for estimating the vertical profiles of wind velocity, temperature and the variances of the vertical and lateral wind speed fluctuations. Selections are available for estimating the mixing height and the surface layer scaling parameters. Only NWS data were input to the processor. This included the hourly Augusta weather observations and the Athens, Georgia (160 km northwest) upper air reports.

In brief, the processor begins its analysis by estimating the surface fluxes and the mixing height. It then constructs estimates of the profiles of wind, wind direction, turbulence and temperature. Further discussion of these computations

and comparisons with tower data are available in Irwin (1985) and Paumier *et al.* (1986).

For estimating the surface scaling parameters, we followed the procedures discussed by Holtslag (1984). The procedures parameterize the surface energy fluxes by cloud cover, time of day, near-surface wind speed, air temperature and surface roughness (assumed to be 0.3 m). Using surface layer similarity relationships, the processor estimates the Monin–Obukhov length, L, and the surface friction velocity, u^*, as continuous variables of the parameterized surface energy fluxes. The daytime convective mixing height estimates are from Garrett's (1981) model. This model employs the rate equations derived by Tennekes (1973) for a zero-order jump model, coupled with a free-entrainment formula based on the work of Deardorff (1974). A procedure outlined by Benkley and Schulman (1979) adjusts the observed shelter temperature for large-scale advection effects, as estimated from the changes in the 700 mb temperatures between upper air reports. Differences in surface temperature at the surface station and the upper air station were handled by adding the potential temperature difference to each of the sounding level temperatures.

The processor estimates the wind velocity profile and the turbulence profiles using the procedures outlined by Holtslag *et al.* (1985). This procedure employs similarity functions as first guess estimates of the vertical profiles of wind speed and turbulence within the mixed layer. The procedures outlined by Van Dop *et al.* (1981) then match the wind profiles below the mixing depth, as estimated using similarity formulas, to the wind profile determined from the twice-daily upper air reports above the mixing depth.

The processor employs data typical as might be used in routine regulatory analyses. We anticipate that the differences seen in observed and simulated meteorological conditions in these data are typical of what might occur in routine analyses. Our intent was not to perform a definitive work of how best to model input uncertainty. The primary question was, are the error distributions skewed, and if so, how might the model input be characterized.

Maximum concentrations from an elevated source occur during convectively unstable conditions when convective eddies are most effective in dispersing the effluent towards the surface. We selected the time periods when conditions are likely to be convectively unstable, hours 1200–1600 local standard time, in the month of June. Data were available at several levels on the meteorological tower. We selected for study the data collected at 300 m since the effluent from sources with stack heights in the range of 75–125 m typically attains a height of 200–600 m.

Therefore comparisons of wind speeds, standard deviations of wind azimuth angle, and standard deviations of wind elevation angle from the processor and from the tower measurements were made. For our purposes, it was sufficient to use graphical techniques to determine whether the error distributions were skewed. We found it convenient to construct cumulative frequency distributions (*cfd*) using subroutine USPRP in the IMSL package, as it provides characterizations for several standard probability density functions. The normal, half-normal, exponential and extreme value distributions were compared with the *cfd* of the residuals (processor minus tower data). The normal, log-normal, half-normal, exponential, Weibull and extreme value distributions were compared with the *cfd* of the ratios (processor estimate divided by tower observation). No one distribution consistently characterized either the residuals or the ratios. The log-normal distribution of the ratios was most often the best representation. Figure A-1 is an example.

APPENDIX B. PLUME RISE AND STACK HEIGHT EFFECTS

The interquartile range, *IQR*, and median of the ratio distributions were found to be primarily functions of the 75th percentile value, *P*75, used to describe the uncertainty associated with the input to the dispersion model. The slight variation of the median values, for different values of plume rise and stack height, is insignificant. Some noticeable improvement can be attained in characterizing the variation of *IQR*, by including the plume rise and stack height into the formulation.

Figure B-1 illustrates the variation of *IQR* as a function of

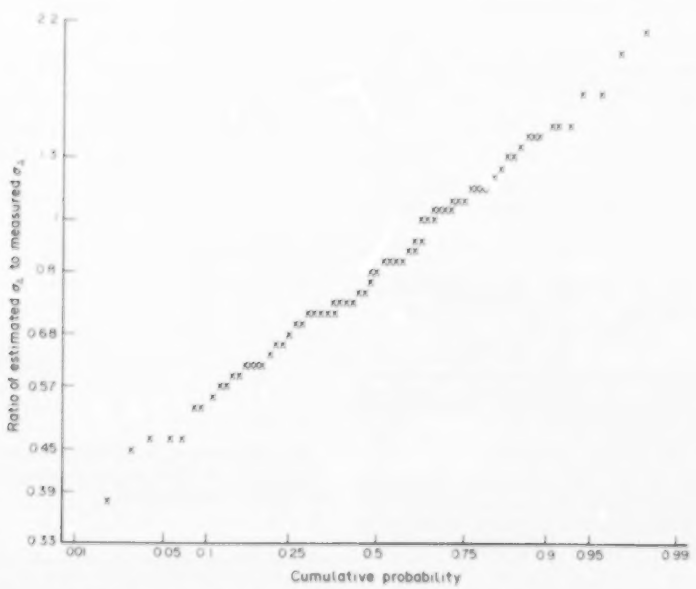

Fig. A-1. Cumulative probability of 15-min ratio of estimated σ_a to measured σ_a at the top level of tower (304 m) for 1500 (local time) for January 1976.

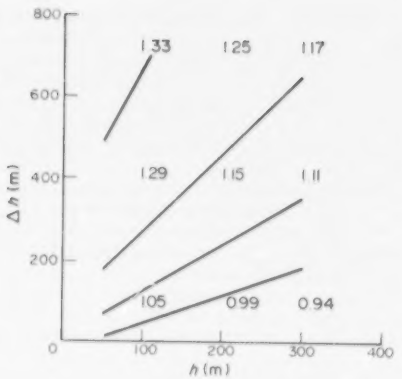

Fig. B-1. Variation of Cm/Cm° interquartile range, IQR, as a function of stack height, h, and plume rise, Δh, for an input 75th percentile, $P75$, of 1.3. Isopleths depicted in figure were generated using Equations (B-1) and (B-2) and the regression results listed in Table B-1.

plume rise, Δh, and stack height, h, for a $P75$-value of 1.3, for the maximum concentration results. Isopleths of IQR appear to fan across the figure, much as the spokes of a wheel, the hub of which is down and to the left of the figure. This variation can be expressed as:

$$IQR = Y/S + T \tag{B-1a}$$

$$Y = \arctan[\Delta h + CDH)/(h + CHS)], \text{ in degrees.} \tag{B-1b}$$

CDH and CHS are the coordinates of the hub. We solved for these coordinates by performing a numerical search, fitting Equation (B-1) by least-squares, locating the position which resulted in a minimum value for the sum squared deviations about the regression line. This search was performed for each $P75$ value. We found the variation in CDH and CHS to be only a weak function of $P75$. Presented in Table B-1 are the average values determined for CDH and CHS.

Given average values for CDH and CHS, we determined S and T by a least squares analysis. The variation of S and T can be characterized using a polynomial as:

$$S = a1 + a2(P75) + a3(P75)^2 \tag{B-2a}$$

$$T = b1 + b2(P75) + b3(P75)^2. \tag{B-2b}$$

Table B-1 lists the values determined for S and T in Equation (B-1) and the polynomial constants for Equation (B-2), for the four features.

Table B-1a. Regression results determined using Equation (B-1) and polynomial constants for use with Equation (B-2), for characterizing the variation of IQR as a function of plume rise (Δh), stack height (h) and input uncertainty ($P75$)

	Maximum concentration ratio		
	$P75 = 1.1$	$P75 = 1.2$	$P75 = 1.3$
S	401.61	212.77	134.95
T	0.27	0.52	0.74
$R1$	0.02	0.18	0.59
$R2$	1.54	5.62	14.09
	$CHS(m) = 85$	$a1 = 9857.41$	$b1 = -4.760$
	$CDH(m) = 90$	$a2 = -14739.44$	$b2 = 6.463$
		$a3 = 5585.13$	$b3 = -1.719$

	Distance to maximum concentration ratio		
	$P75 = 1.1$	$P75 = 1.2$	$P75 = 1.3$
S	653.59	317.46	210.08
T	0.29	0.56	0.83
$R1$	0.02	0.04	0.07
$R2$	1.32	5.53	12.58
	$CHS(m) = -185$	$a1 = 19356.98$	$b1 = -2.447$
	$CDH(m) = 260$	$a2 = -29521.40$	$b2 = 2.306$
		$a3 = 11379.26$	$b3 = 0.164$

$R1$ is the sum squared deviations about (B-1) and $R2$ is the sum squared deviations assuming no Δh and h dependence. ($R1$ and $R2$ are multiplied by 100.)

Table B-1b. Regression results determined using Equation (B-1) and polynomial constants for use with Equation (B-2), for characterizing the variation of IQR as a function of plume rise (Δh), stack height (h) and input uncertainty ($P75$)

	Area of 50% isopleth ratio		
	$P75 = 1.1$	$P75 = 1.2$	$P75 = 1.3$
S	507.61	268.82	184.84
T	0.60	1.11	1.57
$R1$	0.09	0.30	0.57
$R2$	2.37	8.42	17.76
	$CHS(m) = -165$	$a1 = 13353.77$	$b1 = -8.378$
	$CDH(m) = 225$	$a2 = -20192.38$	$b2 = 10.949$
		$a3 = 7740.34$	$b3 = -2.537$

	Area of 25% isopleth ratio		
	$P75 = 1.1$	$P75 = 1.2$	$P75 = 1.3$
S	1515.15	751.88	476.19
T	0.50	0.95	1.39
$R1$	0.20	0.57	0.86
$R2$	1.08	4.17	9.78
	$CHS(m) = -375$	$a1 = 42900.30$	$b1 = -5.700$
	$CDH(m) = 5$	$a2 = -65027.28$	$b2 = 6.639$
		$a3 = 24917.85$	$b3 = -0.913$

$R1$ is the sum squared deviations about (B-1) and $R2$ is the sum squared deviations assuming no Δh and h dependence. ($R1$ and $R2$ are multiplied by 100.)

APPENDIX C. RELATING INPUT ERROR BOUND INTERVALS TO MODEL OUTPUT ERROR BOUND INTERVALS

If we assume a distribution to be log-normal, then given M, the median, and IQR, the interquartile range; the geometric standard deviation, s_g, can be defined as:

$$s_g = ([R + (R^2 + 4)^{1/2}]/2)^{1/p}, \tag{C-1}$$

where $R = IQR/M$ and $p = 0.67419$. For a log-normal distribution, the median equals the geometric mean, m_g. Given m_g and s_g, the log-normal distribution is completely defined. For instance, any value in the distribution, $C(Z)$, is given as:

$$C(Z) = m_g s_g^Z, \tag{C-2}$$

where $Z = (C(Z) - m_g)/s_g$. The quartiles of the distribution occur at plus and minus $Z = 0.67419$; 95% of the values are between plus and minus $Z = 1.96$. A general solution for the error bounds of a log-normal distribution, EB, is:

$$EB = M\{[R + [R^2 + 4]^{1/2}/2]^{1/p}\}^q. \tag{C-3}$$

With $+q$, we specify the upper bound, while $-q$ yields the lower bound. M, R and p are as defined before. Only, we solve for IQR and M, given $R = IQR/M$, using Equations (1a) and (1b) in the text and the regression results presented in Table 3. The value selected for q determines the error bound interval. For instance, $q = p$ yields 50% error bounds, and $q = 1.96$ yields 95% error bounds. If 50% error bounds were constructed, one would expect there is a 50% chance for a given realization's result to reside within the interval between the upper and lower error bound.

As an example, we set $q = p$, and assume that 50% of the input are within a factor of 1.2 ($P75 = 1.2$). Using (1a) and the constants given in Table 3, we find $EB = 1.45$, for the distribution of Cm/Cm°. Stated another way, if the input $P75$-value is 1.2, then 50% of the maximum concentration values are within a factor of 1.45 and 0.69 of the 'error-free' estimate of Cm°.

If we can assume that the input parameter ratios are log-normal and that the distribution of model output ratios are also log-normal, then upper and lower error bounds having different probabilities can be determined as follows.

I. *Basic relationships*

(1) $C(Z) = m_g s_g^z$ Equation (C-2) above.
(2) When $Z = +p$, $C(Z) = P75$
$p = 0.67419$
and when $Z = -p$, $C(Z) = P25$.
(3) $IQR = C(p) - C(-p) = P75 - (1/P75)$
only if $m_g = 1$ is the last relationship true.

II. *Defining and tracking error bounds for the input error distributions*

(1) We assume $m_g = 1$.
(2) We assume $P75$ is a given.
(3) Then from (I1) and (I2) we see that,
$C(p) = m_g s_g^p$ and $C(p) = P75$; $m_g = 1$
so $s_g^p = P75$
$s_g = (P75)^{1/p}$.
(4) EB^i Interval = Upper EB^i − Lower EB^i
$= C(Z) - C(-Z)$
$= [(P75)^{1/p}]^Z - [(P75)^{1/p}]^{-Z}$
^{i} superscript indicates input EB value.
If $Z = p$ then $EB^i =$ the interquartile range $= IQR$
and if $Z = 1.96$ then $EB^i = 95\%$ error bounds.

III. *Defining and tracking error bounds for the output error distributions, given the input* P75*-values*

(1) From Equations (1a) and (1b) in the text and the tabled constants in Table 3, we can define the IQR and M of the output error distributions.
(2) Using (C-1) we then define the output s_g.
(3) EB° Interval = Upper EB° − Lower EB°
$= C'(Z) - C'(-Z)$
$^{\circ}$ superscript indicates output EB value
where $C' = M s_g^Z$ and
$s_g = ([R + (R^2 + 4)^{1/2}]/2)^{1/p}$
$R = IQR/M$ define IQR and M using Equations (1a) and (1b) plus Table 3.

NOTICE

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

This article documents research that was conducted under the auspices of the U.S. Environmental Protection Agency (USEPA). As an arm of the government, the USEPA has a paid up, non-exclusive, irrevocable, worldwide license to reproduce, prepare derivative works, and distribute to the public all of the material contained herein, and to grant others acting on its behalf the right to do so. The USEPA grants to the National Technical Information Service, Department of Commerce, the right to act as its agent in the distribution of this material.